

0590
0328

#P



ENTERED

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,739A

DATE: 04/01/2003

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

4 <110> APPLICANT: Callen, Walter
 5 Richardson, Toby
 6 Frey, Gerhard
 7 Miller, Carl
 8 Kazaoka, Martin
 9 Short, Jay
 10 Mathur, Eric
 12 <120> TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
 13 AND METHODS OF USE THEREOF
 15 <130> FILE REFERENCE: 09010-107001
 17 <140> CURRENT APPLICATION NUMBER: 10/081,739A
 18 <141> CURRENT FILING DATE: 2002-02-21
 20 <150> PRIOR APPLICATION NUMBER: 60/270,495
 21 <151> PRIOR FILING DATE: 2001-02-21
 23 <150> PRIOR APPLICATION NUMBER: 60/270,496
 24 <151> PRIOR FILING DATE: 2001-02-21
 26 <150> PRIOR APPLICATION NUMBER: 60/291,122
 27 <151> PRIOR FILING DATE: 2001-05-14
 29 <160> NUMBER OF SEQ ID NOS: 69
 31 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 33 <210> SEQ ID NO: 1
 34 <211> LENGTH: 1311
 35 <212> TYPE: DNA
 36 <213> ORGANISM: Artificial Sequence
 38 <220> FEATURE:
 39 <223> OTHER INFORMATION: Synthetically generated
 41 <400> SEQUENCE: 1
 42 atggccaagt actccgagct ggaaaaggc ggggtcataa tgcaggcgtt ctactggac 60
 43 gtgccttcag gaggaatatg gtggacaca atacggcaga agataccgga gtggtagat 120
 44 gccggaatct ccgcaatatg gattccccg gcgagcaagg gcatggcgcc cgccatttcg 180
 45 atgggctacg acccctacga cttcttgac ctcggtgagt acgaccagaa gggAACGGTA 240
 46 gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacccgc ccacgcctat 300
 47 ggcatttgcgtaatggccgatatacgatccggccgcggccgtgtatggatggatggatgg 360
 48 aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgctc gggtaataac 420
 49 acggccaaact acctcgactt ccacccgaac gagctccatg cggcgattc cgAACACATT 480
 50 ggaggctatccgacatatg ccacgacaag agctgggacc agtactggct ctggccagc 540
 51 caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac 600
 52 gtcagggtct atgctccctg ggtcgtaag gactggctga actgggtgggg aggctggcg 660
 53 gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt 720
 54 gccaagggtct ttgacttcgc cctctactac aagatggatg aggccttta caacaaaaac 780
 55 attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgaccggc 840
 56 aaggccgtaa cctttgttagc aaaccacgac accgatataa tctggAACAA gtatccagcc 900
 57 tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,739A

DATE: 04/01/2003

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

58 tggctcaaca aggataagct caagaacctc atctggatac atgagaacct cgccggagga 1020
 59 agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
 60 gacaagccgg ggcttataac ctacatcaa cttagtgcga gcaaggccgg aagggtgggtt 1140
 61 tatgtgccga agttcgccgg cgcgtgcattc cacgagtata ctggtaacct cggaggctgg 1200
 62 gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
 63 gccaacgggc agtatggcta ctccgtgtgg agctactgcg ggggtgggctg a 1311
 65 <210> SEQ ID NO: 2
 66 <211> LENGTH: 436
 67 <212> TYPE: PRT
 68 <213> ORGANISM: Artificial Sequence
 70 <220> FEATURE:
 71 <223> OTHER INFORMATION: Synthetically engineered
 73 <400> SEQUENCE: 2
 74 Met Ala Lys Tyr Ser Glu Leu Glu Lys Gly Gly Val Ile Met Gln Ala
 75 1 5 10 15
 76 Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
 77 20 25 30
 78 Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
 79 35 40 45
 80 Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
 81 50 55 60
 82 Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
 83 65 70 75 80
 84 Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
 85 85 90 95
 86 Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
 87 100 105 110
 88 Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
 89 115 120 125
 90 Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
 91 130 135 140
 92 Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
 93 145 150 155 160
 94 Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
 95 165 170 175
 96 Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
 97 180 185 190
 98 Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
 99 195 200 205
 100 Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
 101 210 215 220
 102 Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
 103 225 230 235 240
 104 Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
 105 245 250 255
 106 Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
 107 260 265 270
 108 Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
 109 275 280 285

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,739A

DATE: 04/01/2003

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

110 His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
 111 290 295 300
 112 Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
 113 305 310 315 320
 114 Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Glu Asn
 115 325 330 335
 116 Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
 117 340 345 350
 118 Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
 119 355 360 365
 120 Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
 121 370 375 380
 122 Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
 123 385 390 395 400
 124 Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
 125 405 410 415
 126 Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
 127 420 425 430
 128 Cys Gly Val Gly
 129 435
 131 <210> SEQ ID NO: 3
 132 <211> LENGTH: 1419
 133 <212> TYPE: DNA
 134 <213> ORGANISM: Unknown
 136 <220> FEATURE:
 137 <223> OTHER INFORMATION: Obtained from an environmental sample
 139 <400> SEQUENCE: 3
 140 atgttcctgc tcgcgtttt gctcaactgcc tcgcgtttct gcccaacagg acagcccgcc 60
 141 aaggctgccc caccgttaa cggcaccatg atgcagtatt ttgaatggta ctggccggat 120
 142 gatggcacgt tatggaccaa agtggccaat gaagccaaca acttatccag cttggcatc 180
 143 accgctctt ggctggcccg cgcttacaaa ggaacaagcc gcagcgcacgt agggtacgga 240
 144 gtatacgact tggatgaccc cgccgaaattc aatcaaaaag ggaccgtccg cacaataac 300
 145 ggaacaaaag ctcaatatct tcaagccatt caagccccc acggcgttgg aatgcaagtg 360
 146 tacggcgatg tcgtgttcga ccataaaggc ggcgttgacg gcacggaatg ggtggacgccc 420
 147 gtcgaagtcg atccgtccga ccgcacccaa gaaatctcg gcacctatca aatccaagca 480
 148 tggacgaaat ttgatttcc cggcgcccc aacacctact ccagctttaa gtggcgctgg 540
 149 taccattttt acggcggttga ttgggacgaa agccgaaaat tgagccgcat ttacaaattc 600
 150 cgcggcatcg gcaaagcgtg ggattggaa gtagacacgg aaaacggaaa ctatgactac 660
 151 ttaatgtatg ccgaccttga tatggatcat cccgaagtctg tgaccgagct gaaaaactgg 720
 152 gggaaatggt atgtcaacac aacgaacatt gatgggttcc ggcttgcgtc cgtcaagcat 780
 153 attaagttca gttttttcc tgattgggttgc tcgtatgtgc gttctcagac tggcaagccg 840
 154 ctatttaccg tcgggaaata ttggagctat gacatcaaca agttgcacaa ttacattacg 900
 155 aaaacagacg gaacgatgtc ttgtttgtat gccccgttac acaacaaatt ttataccgt 960
 156 tccaaatca gggcgcatg tgatatgcgc acgttaatga ccaatactct catgaaagat 1020
 157 caaccgacat tggcggtcac cttcggttgc aatcatgaca ccgaacccgg ccaagcgctg 1080
 158 cagtcatggg tcgacccatg gttcaaacccg ttggcttacg cctttatct aactcggcag 1140
 159 gaaggatacc cgtgcgtctt ttatggtgc tattatggca ttccacaata taacattcct 1200
 160 tcgctgaaaa gcaaaatcga tccgctcctc atcgcgca gggattatgc ttacgaaacg 1260
 161 caacatgatt atcttgatca ctccgacatc atcggttgcg caagggaaagg ggtcactgaa 1320

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,739A

DATE: 04/01/2003

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

162 aaaccaggat ccgggctggc cgcaactgatc accgatgggc cgggaggaag caaatggatg 1380
 163 tactgttggc aaacaacacg ctggaaaagt gttctatga 1419
 165 <210> SEQ ID NO: 4
 166 <211> LENGTH: 1539
 167 <212> TYPE: DNA
 168 <213> ORGANISM: Unknown
 170 <220> FEATURE:
 171 <223> OTHER INFORMATION: Obtained from an environmental sample
 173 <400> SEQUENCE: 4
 174 atgaaacaac aaaaacggct ttacgcccga ttgctgacgc tggttatttgc gctcatcttc 60
 175 ttgctgcctc attctgcagc agccggcgca aatctaattg ggacgctgat gcagtatttt 120
 176 gaatggtaca tgcccaatga cggccaacat tggaaagcgct tgcaaaacga ctcggcatat 180
 177 ttggctgaac acggtattac tgccgtctgg attccccgg catataaggg aacgagccaa 240
 178 gcggatgtgg gctacggtgc ttacgaccct tatgatttag gggagttca tcaaaaaggg 300
 179 acggttcggg caaaatgtacgg cacaaaagga gagctgcaat ctgcgatcaa aagtcttcat 360
 180 tcccgcgaca ttaacgttta cggggatgtg gtcataacc acaaaggcgg cgctgatgctg 420
 181 accgaagatg taaccgcggt tgaagtcgat cccgctgacc gcaaccgcgt aatttcagga 480
 182 gaacaccgaa ttaaagcctg gacacattt cattttccgg ggcgcggcag cacatacagc 540
 183 gattttaaat ggcattggta ccattttgac ggaaccgatt gggacgagtc cggaaagctg 600
 184 aaccgcattt ataagttca aggaaaggct tgggatttggg aagttccaa tgaaaacggc 660
 185 aactatgatt atttgatgtt tgccgacatc gattatgacc atcctgatgt cgccagcagaa 720
 186 attaagagat ggggcacttgc gtatgccaat gaactgcaat tggacgggtt ccgtcttgat 780
 187 gctgtcaaac acattaaatt ttcttttttgc cgggatttggg ttaatcatgt cagggaaaaaa 840
 188 acggggaaagg aaatgtttac ggtagctgaa tattggcaga atgacttggg cgcgcggaa 900
 189 aactatttga acaaaaacaaa tttaatcat tcagtgttttgc acgtgccgct tcattatcag 960
 190 ttccatgctg catcgacaca gggaggcggc tatgatatga ggaaatttgc gaacggatcg 1020
 191 gtcgtttcca agcatccgtt gaaagcggtt acatttgcgataaaccatga tacacagccg 1080
 192 gggcaatcgc ttgagtcgac tgtccaaaca tggtttaagc cgcttgctta cgctttcatt 1140
 193 ctcacaaggaa aatctggata ccctcaggat ttctacgggg atatgtacgg gacgaaagga 1200
 194 gactcccagc gcgaaattcc tgccttggaa cacaaaatttgc aaccgatctt aaaagcgaga 1260
 195 aaacagtatg cgtacggagc acagcatgtat tatttcgacc accatgacat tggcggctgg 1320
 196 acaaggaaag ggcacagctc ggttgc当地 tcaggtttgg cggcattaat aacagacgg 1380
 197 cccggggggg caaagcgaat gtatgtggc cggccaaacg ccggtgagac atggcatgac 1440
 198 attaccggaa accgttgcga gccgggttgc atcaatttgc aaggctgggg agagtttac 1500
 199 gtaaacggcg ggtcggttgc aatttatgtt caaagatag 1539
 201 <210> SEQ ID NO: 5
 202 <211> LENGTH: 1395
 203 <212> TYPE: DNA
 204 <213> ORGANISM: Unknown
 206 <220> FEATURE:
 207 <223> OTHER INFORMATION: Obtained from an environmental sample
 209 <400> SEQUENCE: 5
 210 gtgggtcaca tgaagttgaa gtaccttgc ttatgtttgt tggctgtggc ttgcataaggc 60
 211 ctactctcga ctccagtggg tgctgccaag tactccgaac tcgaagaggg cgggttata 120
 212 atgcaggccct tctactggta tggtttccgg gggggatct ggtgggacac cataagacag 180
 213 aaaatcccg agtggatcgatc cgctggatc tcggcgatat ggattccctcc agctagcaaa 240
 214 gggatggcg gtttttattc catgggctac gatccctacg atttcttgc cctcggcgag 300
 215 tactatcaga agggaaacagt tgagacgcgc ttccggctcaa aggaggaact ggtgaacatg 360
 216 ataaacacccg cacactccta tggcataaaag gtgatagcgg acatagtcataa aaccaccgc 420

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,739A

DATE: 04/01/2003

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

217	gccgggtggag	accttggatg	gaaccgcctt	gtaaacaact	atacttggac	agacttctcc	480
218	aaggtcgcct	ccggtaaata	cacggccaac	taccccttgact	tccaccggaaa	cgagggtcaag	540
219	tgctgcgatg	agggtacatt	tggtgacttt	ccggacatcg	cccacgagaa	gagctgggat	600
220	cagtactggc	tctgggcaag	caatgagagc	tacggccat	atctccggag	catagggatc	660
221	gatgcattggc	gtttcgacta	cgtcaaagggt	tacggagcg	gggttggtaa	tgactggctc	720
222	agctgggtgg	gaggctgggc	cggtggagag	tactggaca	cgaacgttga	tgcactcctt	780
223	aactgggcatt	acgacagcgg	tgccaaaggtc	tttgacttcc	cgctctacta	caagatggac	840
224	gaagcctttg	acaacaccaa	catccccct	ttggtttacg	ccctccagaa	cgaggaaaca	900
225	gtcgtttccc	gcatccctt	caaggcagta	actttcggtt	ccaaccacga	tacagatata	960
226	atctggaaaca	agtatccggc	ttatgcgttc	atccttacat	atgagggaca	gcctgttata	1020
227	ttttaccgcg	actacgagga	gtggctcaac	aaggataagc	ttaacaacct	tatctggata	1080
228	cacgagcacc	ttgcccggagg	aagtaccaag	atcctctact	acgataacga	ttagctataa	1140
229	ttcatgaggg	agggctacgg	gagcaagccg	ggcctcataa	cctacataaa	cctcgaaac	1200
230	gactgggccc	agcgctgggt	gaacgtccgc	tcaaagtttgc	ccggctcacac	aatccatgaa	1260
231	tacacaggca	atctcggttgc	ctgggttgc	aggtgggttgc	agtacgtatgg	atgggttaaa	1320
232	ctgacggcac	ctcctcatga	tccagccaa	ggatattacg	gctactcagt	ctggagctac	1380
233	gcaggcgtcg	gatga					1395
235	<210>	SEQ ID NO:	6				
236	<211>	LENGTH:	1386				
237	<212>	TYPE:	DNA				
238	<213>	ORGANISM:	Unknown				
240	<220>	FEATURE:					
241	<223>	OTHER INFORMATION:	Obtained from an environmental sample				
243	<400>	SEQUENCE:	6				
244	atgaagaagt	ttgtcgccct	gttcataacc	atgttttgc	tagtgagcat	ggcagtcgtt	60
245	gcacagccag	ctagcgcgc	aaagtattcc	gagctcgaag	aaggccgcgt	tataatgcag	120
246	gccttctact	gggacgtccc	agggtggagga	atctgggtgg	acaccatcg	gagcaagata	180
247	ccggagtggt	acgaggccgg	aatatccgc	atttggattc	cgccagccag	caaggggatg	240
248	agcggcgtt	actcgatggg	ctacgatccc	tacgatttct	ttgacctcgg	cgagtacaac	300
249	cagaagggaa	ccatcgaaac	gcgctttggc	tctaaacagg	agctcatcaa	tatgataaac	360
250	acggcccatg	cctacggcat	aaaggtcata	gcccacatcg	tcataaaacca	ccgcgcaggc	420
251	ggagacctcg	agtggAACCC	gttcgttggg	gactacacct	ggacggactt	ctcaaagggtg	480
252	gcctcgccca	aatataactgc	caactaccc	gacttccacc	ccaacgaggt	caagtgtgt	540
253	gacgagggca	catttggagg	cttcccagac	atagcccacg	agaagagctg	ggaccagcac	600
254	tggctctggg	cgagcgtatga	gagctacg	gcctaccta	ggagcatcgg	cgttgtatgc	660
255	tggcgctttg	actacgtaa	gggctacgg	gcgtgggtcg	tcaaggactg	gctcaactgg	720
256	ttggccggct	gggcccgttgg	cgagttactgg	gacaccaacg	ttgatgcact	cctcaactgg	780
257	gcctactcga	gcggcgccaa	ggtcttcgac	ttcccgtct	actacaagat	ggatgaggcc	840
258	tttgacaaca	aaaacattcc	agcgtcgtc	tctgcccttc	agaacggcca	gactgttgc	900
259	tcccgcgacc	cgttcaaggc	cgtaacctt	gtagcaaacc	acgacaccga	tataatctgg	960
260	aacaagtacc	ttgcttatgc	tttcatcc	acctacgaag	gccagccgt	catattctac	1020
261	cgcgactacg	aggagtggct	caacaaggac	aggttgaaca	acctcatatg	gatacacgac	1080
262	cacctcgca	gtggaaagcac	gagcatagtc	tactacgaca	gcgacgagat	gatcttcgtg	1140
263	aggaacggct	atggaaagcaa	gcctggcctt	ataacttaca	tcaacctcgg	ctcgagcaag	1200
264	gttggaaagg	gggttatgt	gccgaagttc	gcgggcgcgt	gcatccacga	gtatactgg	1260
265	aacctcgag	gctgggtaga	caagtacgtc	tactcaagcg	gctgggtcta	tctcgaaagct	1320
266	ccagcttacg	accctgccaa	cggcagat	ggctactccg	tgtggagcta	ttgcgggttt	1380
267	gggtga						1386
269	<210>	SEQ ID NO:	7				

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/081,739A

DATE: 04/01/2003

TIME: 13:17:24

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw